

ORIGIN AND TRANSMISSION PATTERNS OF SARS-COV-2 IN LOMBARDY

Prof. Carlo Federico PERNO – Università degli Studi di Milano

Methods

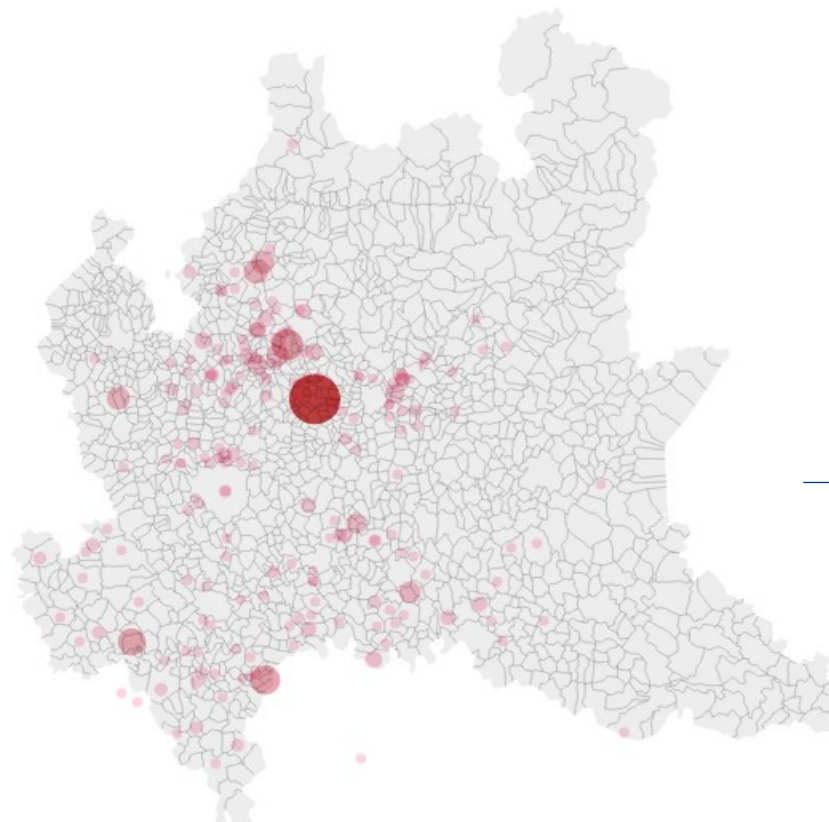
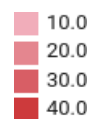
- A retrospective cohort study that included 346 SARS-CoV-2 sequences collected by nasopharyngeal-swabs of adult patients diagnosed for SARS-CoV-2 infection at ASST Grande Ospedale Metropolitano Niguarda (Milan, Italy) and at IRCCS San Matteo Hospital (Pavia, Italy) since February 22
- Patients resided in all 12 provinces of Lombardy and with varying disease symptoms, ranging from mild to severe

Results / 1

- From February 22 through April 4, 2020, nasopharyngeal swabs of a total of 25,082 patients were screened for SARS-CoV-2 infection at two major Hospitals in Lombardy. Of them, 11,445 received a diagnosis of COVID-19
- Whole genome sequencing was performed in a total of 371 samples collected from 371 patients residing in all 12 provinces of Lombardy and with varying disease symptoms, ranging from mild to severe. Twenty-five samples were excluded due to failed amplification (n=9), or poor genomic coverage (<60%, n=16). The final study population thus consisted of 346 patients

Results / 2

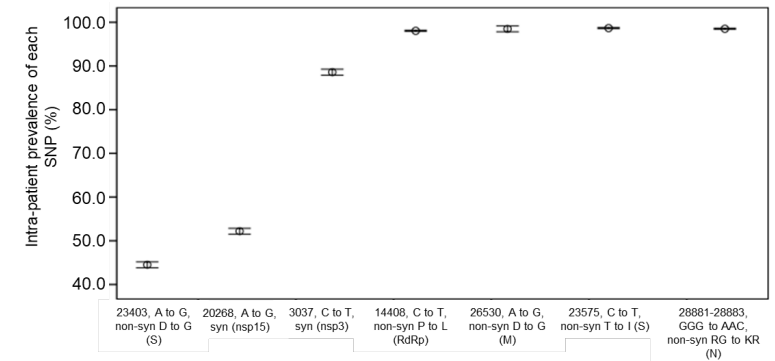
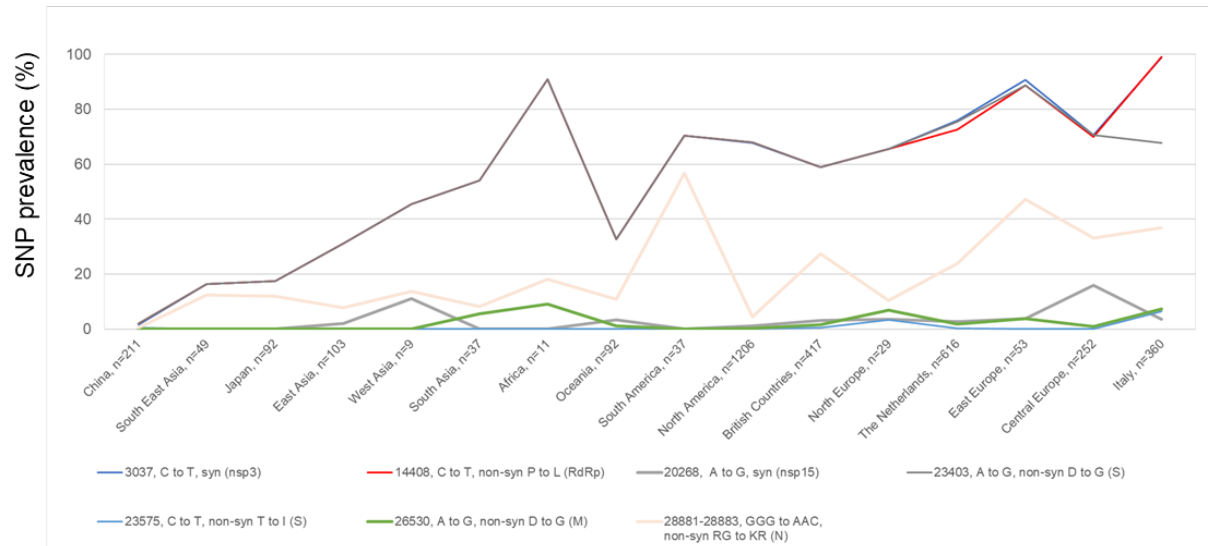
- Geographic distribution of COVID-19 cases among the 12 provinces of Lombardy



Milan:	75
Como:	69
Pavia:	65
Bergamo:	36
Lecco:	32
Lodi:	32
Cremona:	27
Other:	19

Results / 3

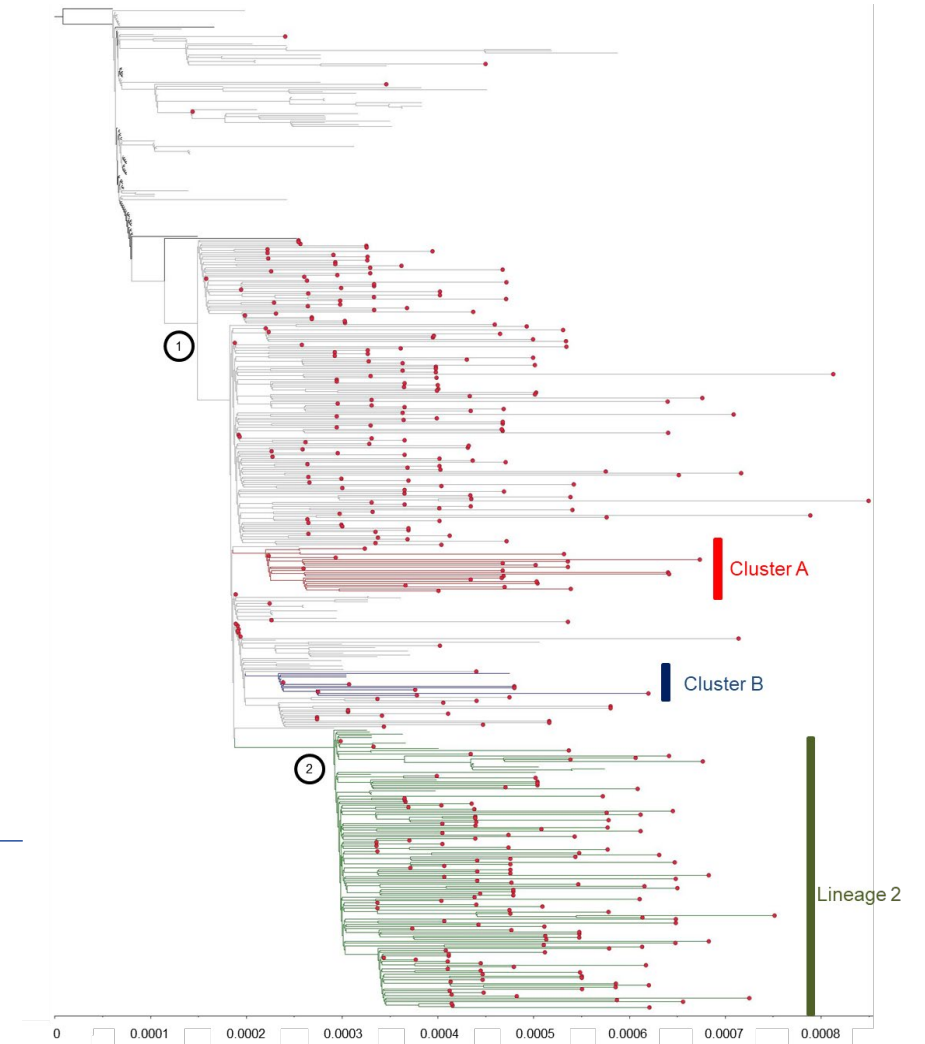
- > Looking at the overall variability of SARS-CoV-2, we found that **only 7 SNPs (2 out of 7 synonymous)** characterized our consensus sequences, **highlighting a good conservation rate of this virus along time**. This conservation rate is confirmed within the spike structural protein, where only 2 mutations, one of them at low prevalence (i.e, C23575T, corresponding to the amino acid variant T671I), were detected
- > None of these mutations have a role in altering pre-existing N-glycosylation sites or in creating new ones, while their role in compromising the vaccine efficacy requires further investigations



Results / 4

- ML tree showed that **most of SARS-CoV-2 sequences from Lombardy** (grey taxa with red dots, 342/346 [98.8%]) **are interspersed within two major viral lineages** (1, 2) containing previously isolated Italian strains, and strains from other European, American and Asian countries (gray branches, no dots)

Gisaid sequences come from Central Italy (n=6), East Europe (n=6), North Europe (n=6), South America (n=6), Africa (n=7), Japan (n=7), Oceania (n=7), West Asia (n=7), North Italy (n=9), South Asia (n=9), Central Europe (n=11), East Asia (n=11), European Low Countries (n=12), South East Asia (n=13), North America (n=18), British Countries (n=26), China (n=36).



Results / 5

- These lineages did not contain viral strains isolated in the first months of the outbreak in China (black branches, no dots); **this let us hypothesize a transmission chain not directly involving China (i.e., the country where the pandemic originated)**
- **Notably, the most closely related viral isolate that clustered outside such Lineages was isolated in Central Europe, in the second half of January**

Results / 6

- Both the SARS-CoV-2 lineages we identified have a quite simultaneous ancestry, dating at least one month before the first diagnosed cases
- **Even if the two lineages co-circulate in overlapping ranges of territories, Lineage 1 was dominant in the South, mostly in Lodi and Cremona, while Lineage 2 predominated in the North of Lombardy, mostly in Bergamo and its adjacent territories ravaged by the virus (like Alzano and Nembro)**
- *The predominance of these lineages in different territories within their concurrent circulation, the lack of a monophyletic signal for Lineage 1, and the detection of two well supported clusters inside it, supports the hypothesis of multiple and dislocated introductions of SARS-CoV-2 with different route of viral transmission*

Discussion

- This study allows the identification of SARS-CoV-2 lineages circulating in the most affected COVID-19 area in that period, representing a huge reserve of genetic information of a virus that became able to a pandemic spread, and caused more than 300,000 deaths in some weeks
- We cannot exclude that this multiple and simultaneous circulation of SARS-CoV-2 strains can have exacerbated the transmissibility potential of the virus and thus create a real viral storm in such high densely populated region

Conclusion

- The characteristics of the virus indicate that it entered by the middle of January (or even before) in various geographic areas of Lombardy (most probably at the same time)
 - Ⓒ This can contribute to the understanding why the virus was so aggressive and difficult to contain
- The very limited variability of the virus (only 7 mutations in a total of about 30,000 genomic positions) suggests that the virus had a minimal evolution during the observation period
 - Ⓒ This tends to exclude a change in the infectivity and/or in the pathogenic potential of the virus, at least in this time frame
 - Ⓒ Therefore, the evidence of a recent decreased aggressivity of the disease has to be searched among reasons not linked to the characteristics of the virus (Earlier diagnoses? Better control of viral spreading? Warmer air temperature? Other reasons?)
 - Ⓒ ***This limited variability of the virus supports the chances of a vaccine to be effective against SARS-CoV-2, provided that it is properly designed, developed and administered***